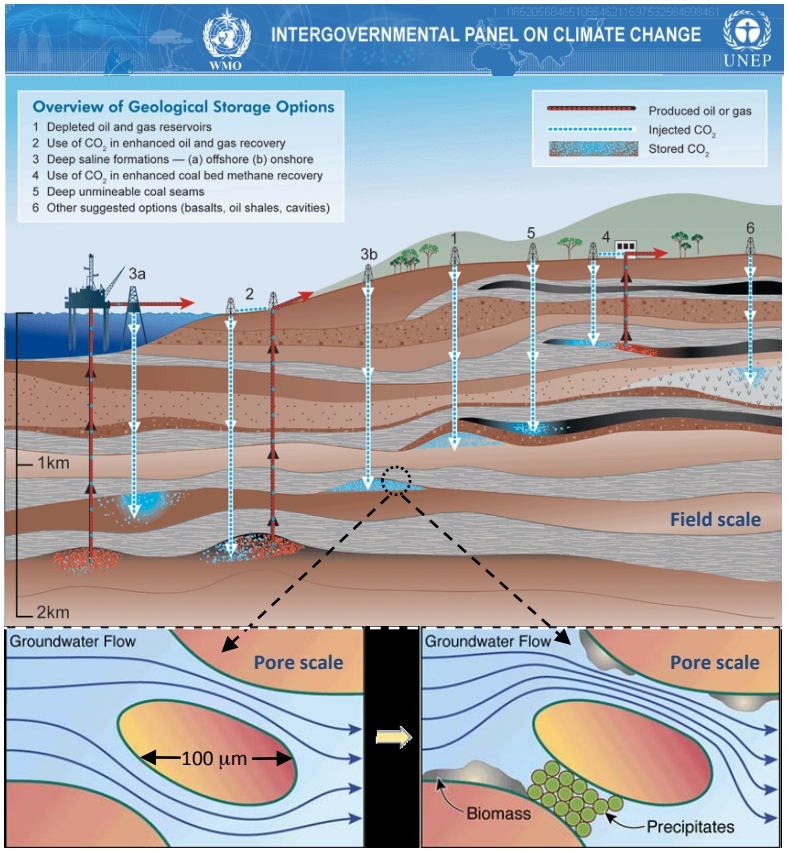
Simulation of Pore Scale Reactive Transport Processes Associated with Carbon Sequestration

Present and Future Computing Requirements for Chombo-Crunch

David Trebotich
Computational Research Division, LBL

Our goal is to enable accurate prediction of the fate of geologically stored CO_2

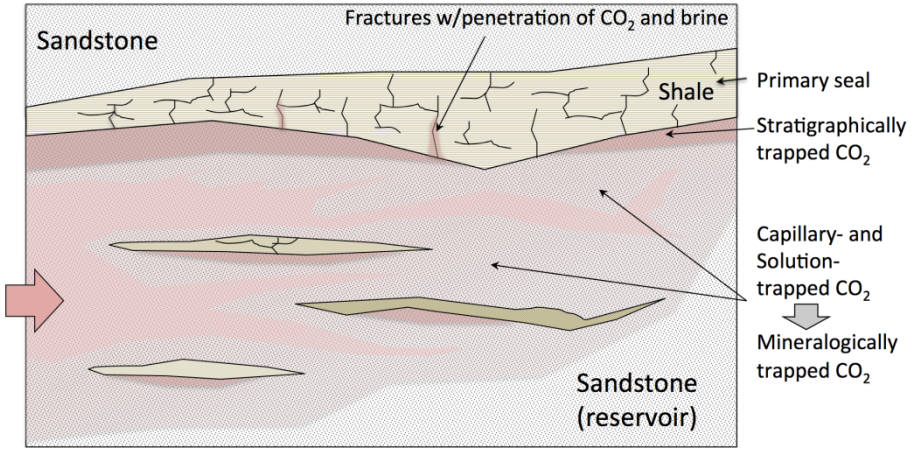
Flow and transport typically simulated at field scale



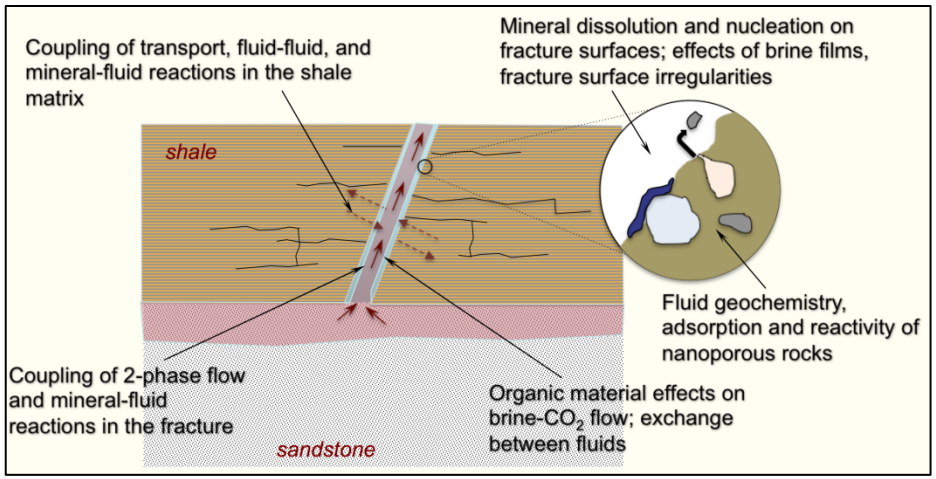
CO_2 trapping mechanisms governed by emergent processes at pore (micro) scale

➔ **Need high resolution pore scale reactive transport model**
➔ **Need methods to upscale pore information to field scale**

We will use hpc to simulate pore scale reactive transport at unprecedented scale



First round of EFRC modeling and simulation effort focused on minerological trapping



Second round will additionally focus on stratigraphic trapping

➔ **Mesoscale modeling challenge**

- **Shales are tight cf. carbonates**
- **Multiphase effects more prominent, triple point problem**

We have developed high performance algorithms and simulation tools in Chombo

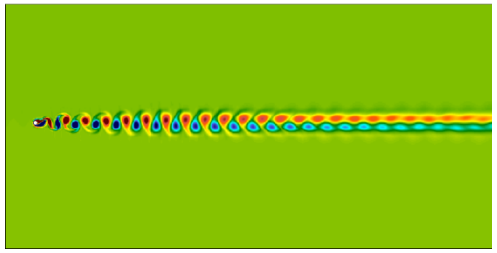
Chombo EBAMRINS application

Incompressible flow and transport solver in complex geometries

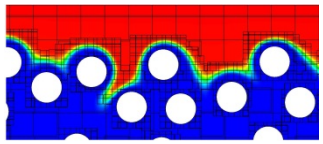
- Embedded boundary (EB) method
  - Finite volume approach
  - Accurate reactive surface area
  - Dynamic local refinement (AMR)
  - AMG elliptic solvers (PETSc)
  - DNS from image data
  - Optimized for high performance
  - hdf5 I/O; VisIt diagnostics
- Consistent approach to interfaces
  - fluid-mineral (precip./dissolution)
  - fluid-fluid (multiphase)

2D flow past cylinder, Re=300, >100d wake

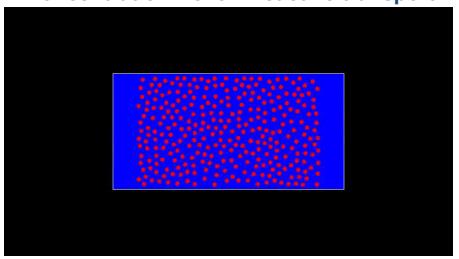
3D flow past cylinder, Re=300, >100d wake



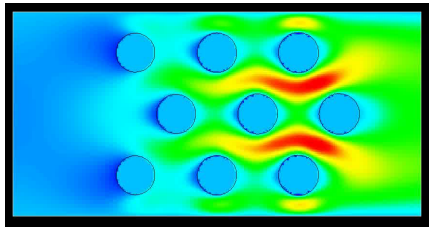
Concentration front with AMR



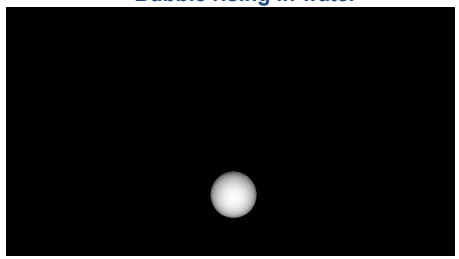
Concentration front in reactive transport



Flow past diminishing cylinders



Bubble rising in water



Chombo-Crunch simulates microscale flow, transport and geochemical reactions

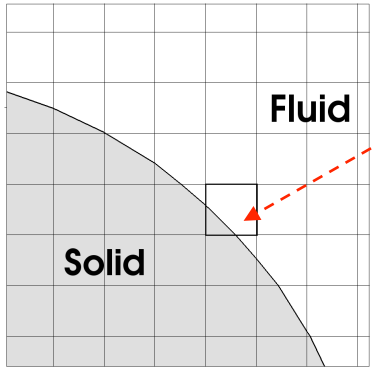
Chombo-Crunch

CFD + single phase multi-component geochemical reactive transport in very complex pore (micro) scale geometries

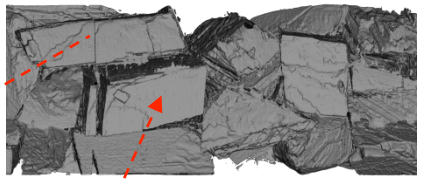
- Adaptive, finite volume methods for advection-diffusion in Chombo
  - Accurate reactive surface area using embedded boundaries
  - Dynamic local refinement (AMR)
  - Scalable (100K processors)
  - DNS from geologic image data

- CrunchFlow geochemistry
  - SNIA operator splitting
  - Point-by-point calculation
  - CFL-limited timestep

Embedded boundary representation



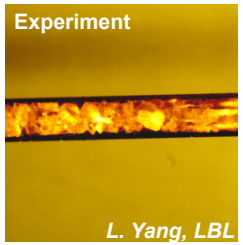
Computational domain for calcite in capillary tube



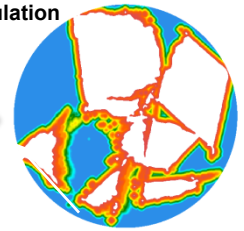
Rate calculated at each water-mineral interface by multiplying by the reactive surface area (RSA)

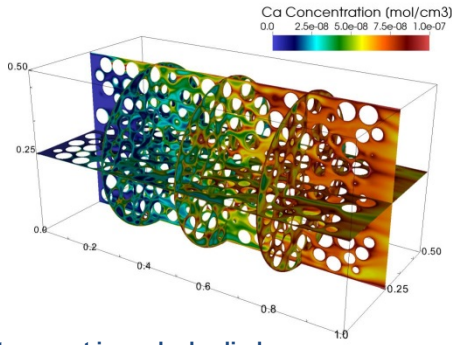
Image data converted to simulation grid using implicit function representation of boundaries

Calcite in capillary tube

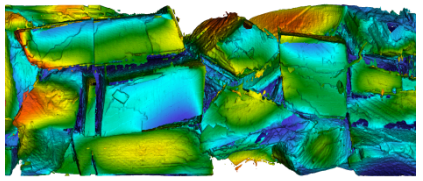


Simulation





Reactive transport in packed cylinder

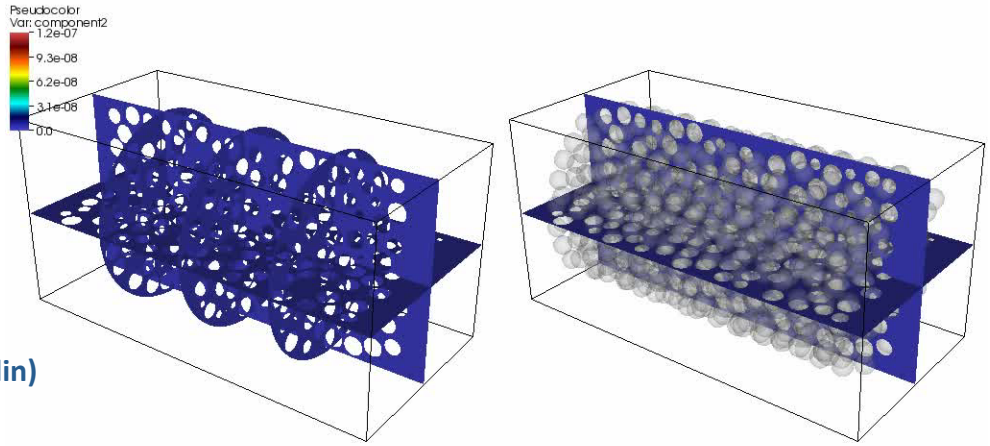


pH on crushed calcite in capillary tube

Chombo-Crunch has been used to model steady-state multicomponent RT

Total calcium in calcite dissolution reaction network

- 1 cm long cylinder
- 1000 spheres, 250 μm radii
- <20 μm grid resolution
- 1024 processor cores (Franklin)



Implicit function representation of geometry

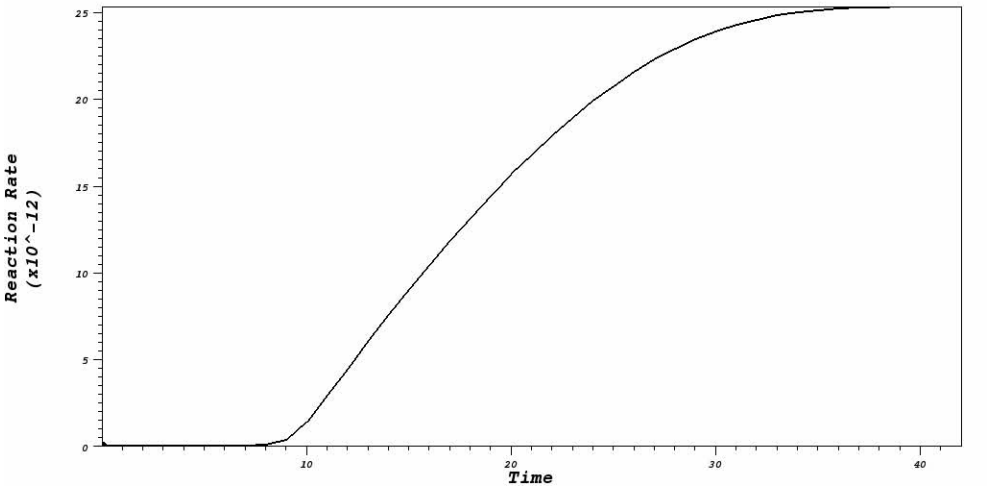
$$\phi(x) = \min_k(|\vec{x} - \vec{x}_k|^2 - r_k^2)$$

$$\vec{x}_k = center\ of\ kth\ sphere$$

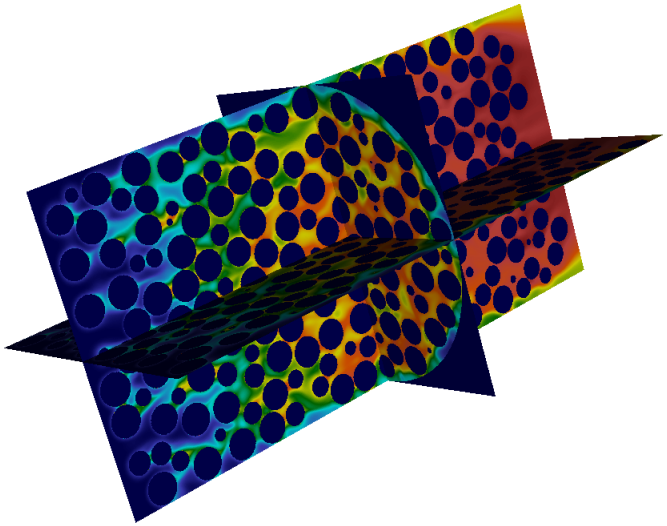
$$r_k = radius\ of\ kth\ sphere$$

$$\vec{x} : \phi(\vec{x}) = 0\ boundary$$

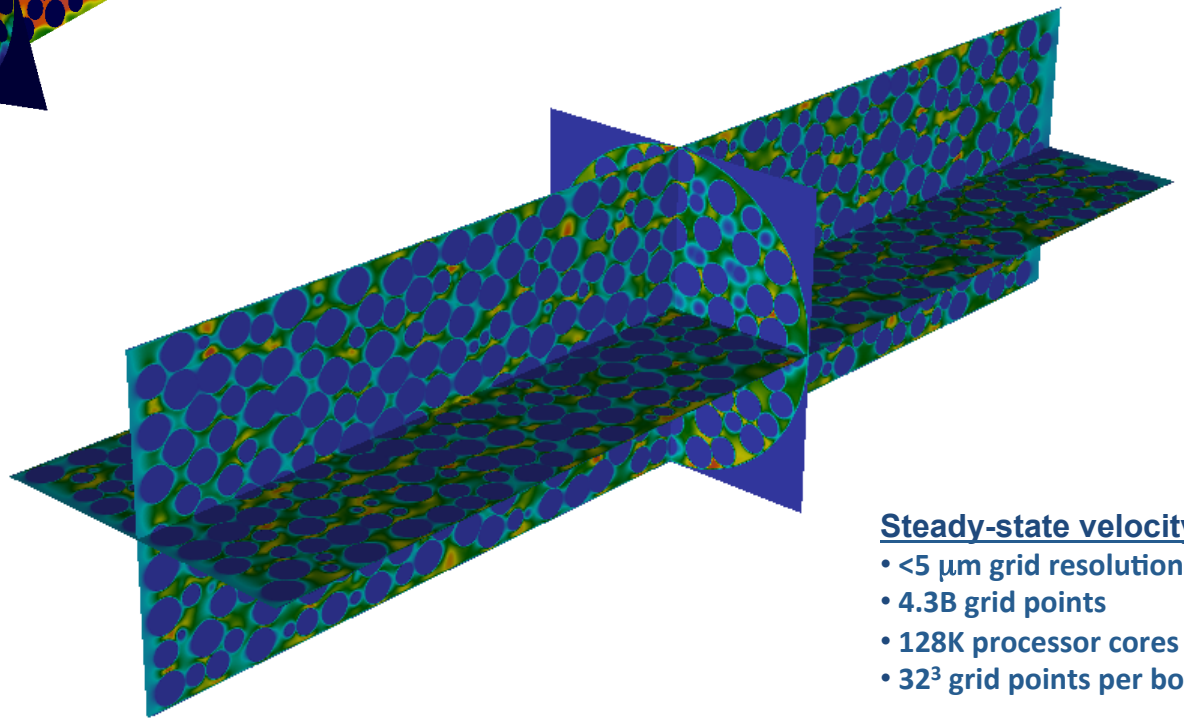
Total reaction rate evolving to steady-state



Chombo-Crunch is now a production code for multi-component reactive transport

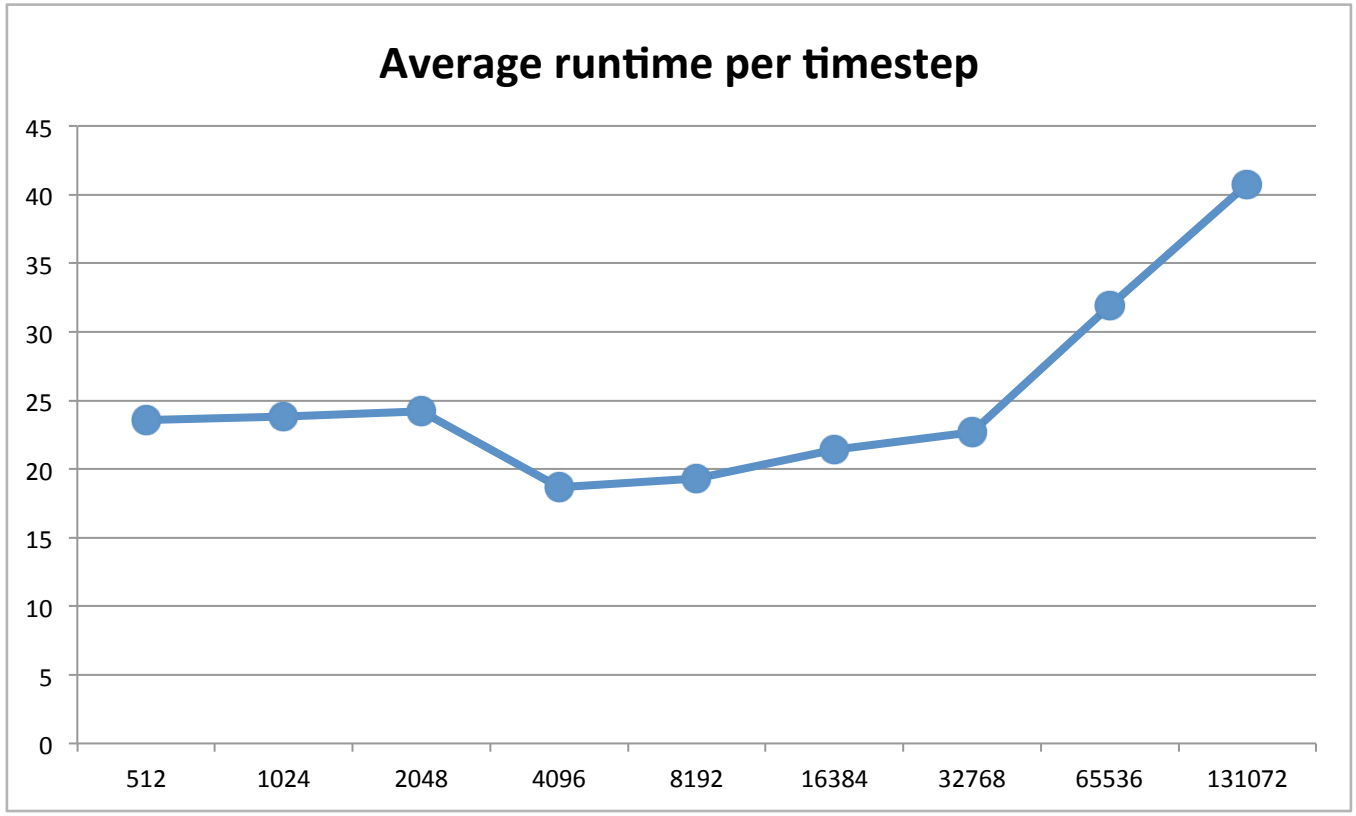


Transient total calcium
- <5 μm grid resolution
- 2.15B grid points
- 64K processor cores (Hopper/Edison)
- 32^3 grid points per box



Steady-state velocity
- <5 μm grid resolution
- 4.3B grid points
- 128K processor cores (Hopper)
- 32^3 grid points per box

Chombo-Crunch scales to 100,000s processor cores on NERSC machines

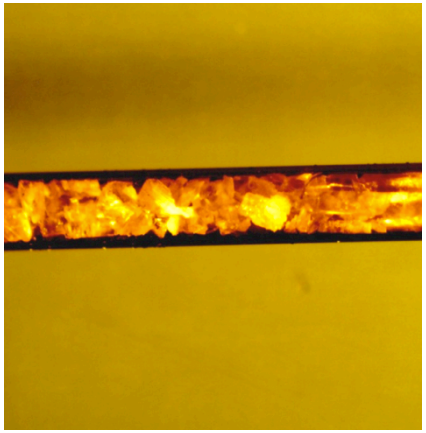


➔ **Chombo-Crunch is memory bandwidth limited**

➔ **Load balancing sweet spot:**

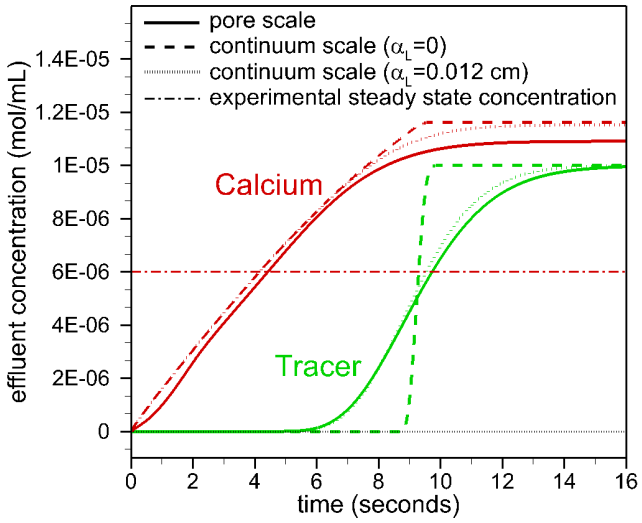
- **32^3 grid points per box, 1 box per core**

Science at scale for CO_2 sequestration 2012-13 NISE award



EFRC-NCGC validation experiment

- Reactive transport in capillary tube packed with calcite
  - 7 mm long x 500 μm diameter
  - 0.899 μm image resolution
  - 9.3 seconds residence time (steady-state)
- Chombo-Crunch simulation
  - 1.6B simulation grid points
  - 1.19 μm resolution
  - 700GB plot files, 600GB checkpoint files
  - 49,152 cores on Hopper Cray XE6
  - 100 time steps per hour
  - 1 time step = 0.000182 seconds real time



➔ Current rate law in reactive transport model overshoots steady-state effluent concentration

➔ Run coarser simulation (2μm) with new rate law, cf. Edison…

Preliminary Chombo-Crunch runs on Cray XC30 outperform XE6 by 2.3x overall run time

Hopper

2 μm simulation resolution

6144 cores for 36 hours

3567 time steps

1 time step = 0.000265 seconds of simulated time

XE6 Performance

- 35.5 seconds per time step
  - 1.69 time steps per minute
  - 100 time steps per hour
- Lustre IO = 1.8% of total run time
  - 1.1% checkpoint files
  - 0.7% plot files
- < 1% initialization overhead

Edison

2 μm simulation resolution

6144 cores for 6 hours

1366 time steps

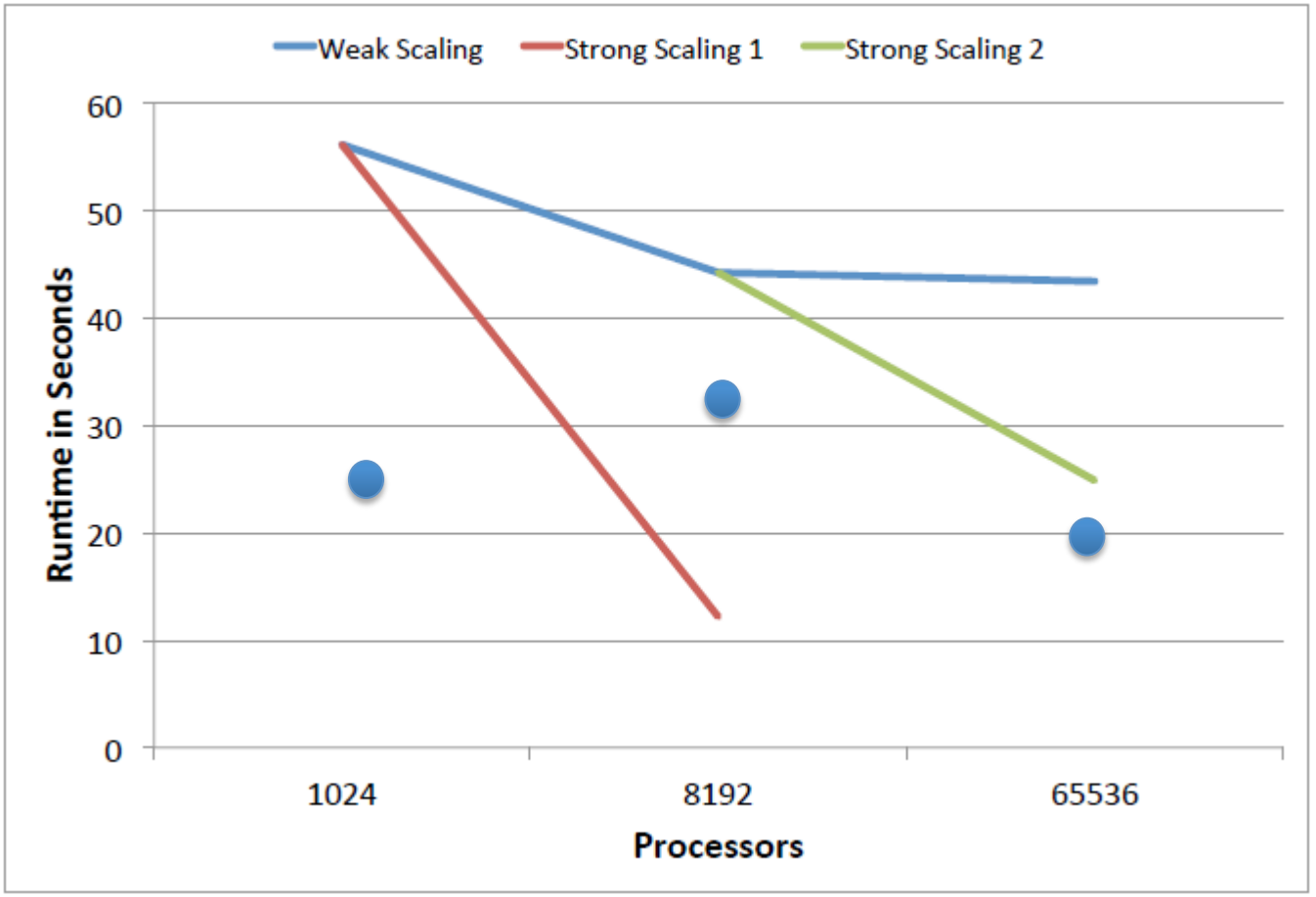
1 time step = 0.000265 seconds of simulated time

XC30 Performance

- 15.3 seconds per time step
  - 3.92 time steps per minute
  - 230.77 time steps per hour
- Lustre IO = 2.3% of total run time
  - 1.8% checkpoint files
  - 0.5% plot files
- 1% initialization overhead

➔ Chombo-Crunch is memory bandwidth limited

Preliminary Chombo-Crunch runs on Cray XC30 outperform XE6 by 2.3x overall run time



Edison Phase 2

The embedded boundary method resolves fine features of calcite grains

Computational domain – 1 micron grid resolution (1.6B grid points)

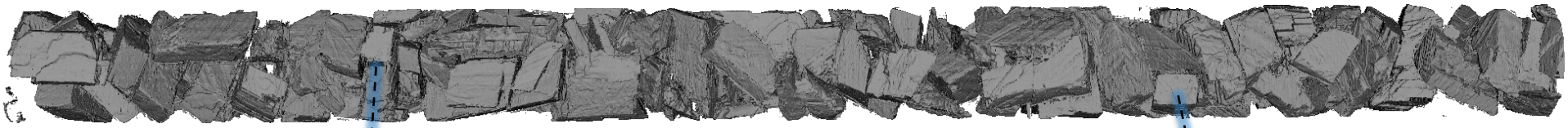
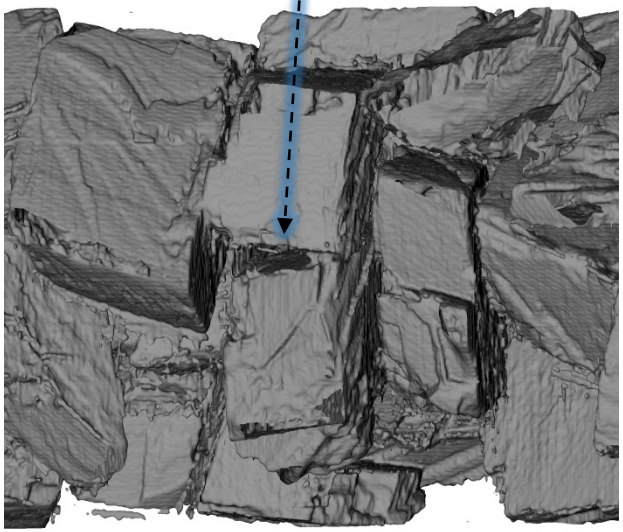
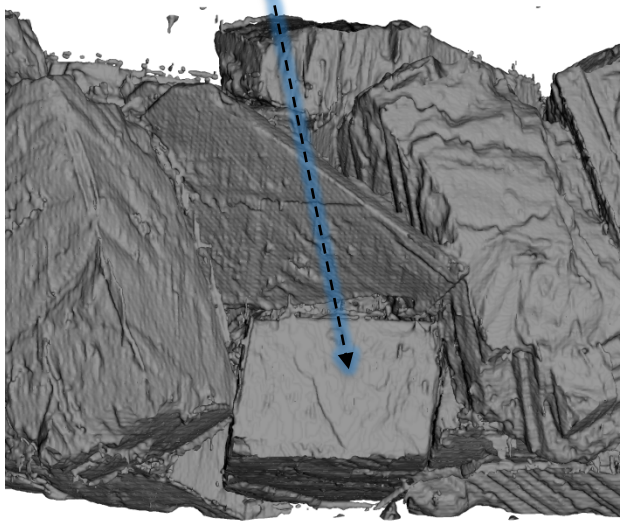


Image to simulation grid using implicit functions

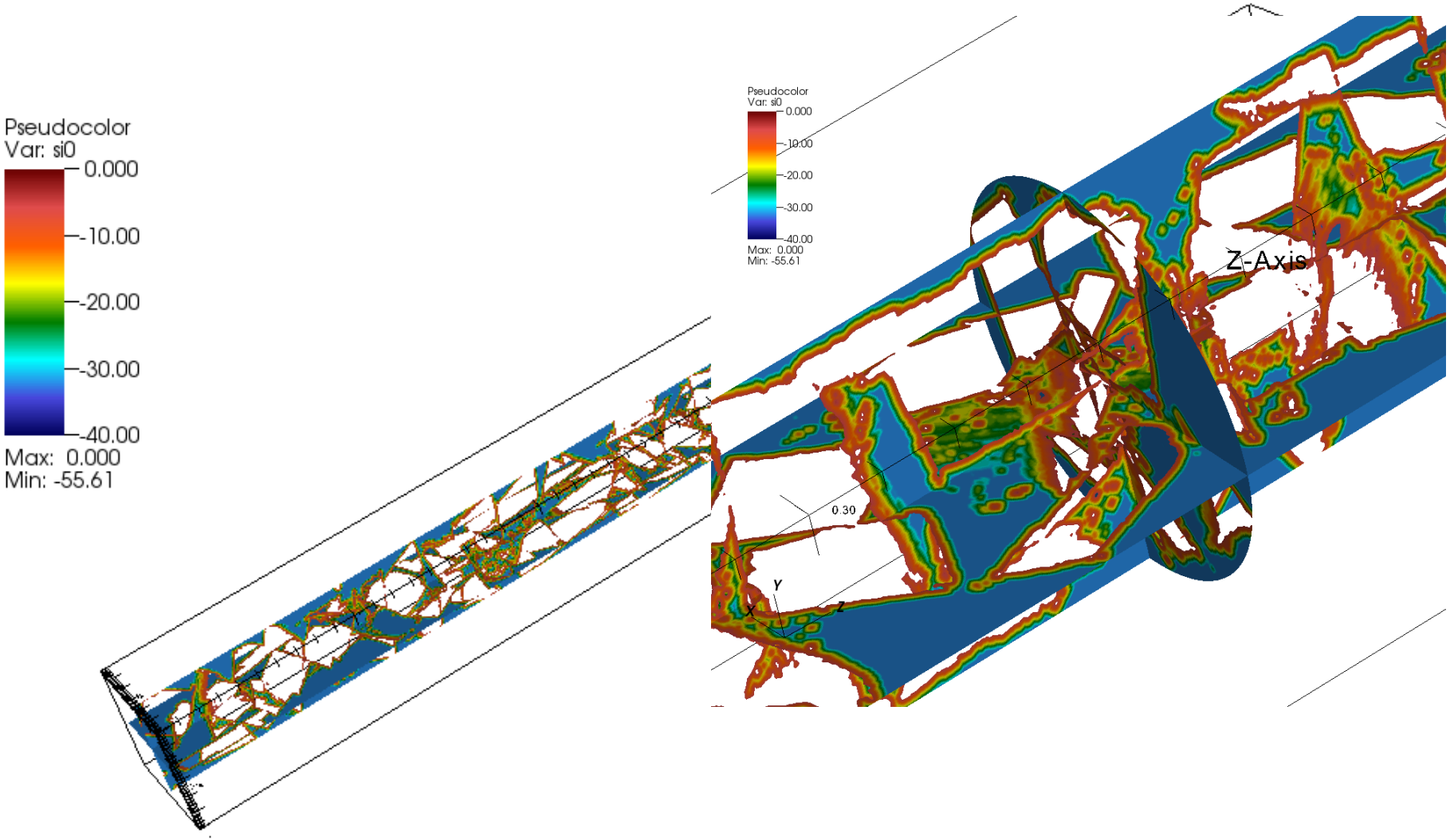




➔ **RSA from EB = 0.00137 m^2/g**

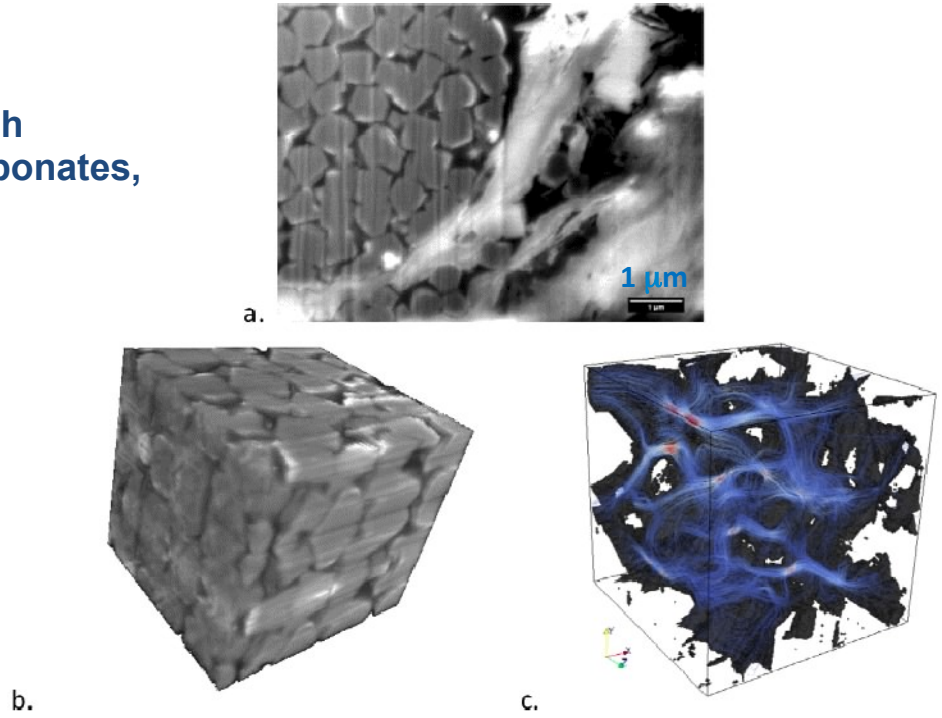
cf. BET = 0.0012 m^2/g (BET error ~ 15%)

High resolution shows microscale features
Diffusion boundary layer



Flow and transport in shales is a mesoscale modeling challenge

Pore space is much tighter than in carbonates, possible fractures



Flow streamlines in a Haynesville shale sample (Dewers et al, 2012), (bottom) streamlines through the porespace in a framboidal pyrite region in New Albany shale (Silin and Kneafsey, 2012)

100 nm resolution floor to be consistent with continuum assumption (Tretheway and Meinhart 2002)

Requirements for EFRC Mesoscale Challenge Reactive Transport in Shales

Year	Problem dimension	Domain size	Grid points	Spatial resolution	Problem time scale	HPC resources	Problem DOF	Data storage per plot file
2013	2D	1m	2B	O(1 μm)	O(hours)	50M hrs, 1PB	25	1TB
	3D	1cm	2B	O(1 μm)	O(minutes)	100M hrs, 1PB	25	1TB
2017	2D	10cm	20B	O(100 nm)	O(hours)	500M hrs, 5PB	5	10TB
	3D	1cm	20B	O(100 nm)	O(minutes)	1B hrs, 5PB	5	10TB

Strategies for new architectures

Does your software have CUDA/OpenCL directives; if yes, are they used, and if not, are there plans for this?

No

Does your software run in production now on Titan using the GPUs?

No

Does your software have OpenMP directives now; if yes, are they used, and if not, are there plans for this?

Port Chombo-Crunch to C4 in next year.

Does your software run in production now on Mira or Sequoia using threading?

No

Is porting to, and optimizing for, the Intel MIC architecture underway or planned?

Yes, have an account on Babbage.

Strategies for new architectures

- Have there been or are there now other funded groups or researchers engaged to help with these activities?
  - We have engaged the ExaHDF5 team led by Prabhat,
  - the PETSc team (Mark Adams et al. through FASTMath)
  - ALCF consultants
  - X-stack projects with Sam Williams and Brian Van Straalen

- What role should NERSC play in the transition to these architectures?
  - Tell us what hardware you are considering and then we can tell you how we will use it. (e.g., BB)

- What role should DOE and ASCR play in the transition to these architectures?
  - DOE and ASCR need to engage the algorithm and simulation software developers for guidance on future directions in hardware. For example, I would be set beyond 2017 with a machine like Edison that has 4 or more times the current core count.

What is the effect of heterogeneous pore structure on average reaction rates?

Porosity = 0.762

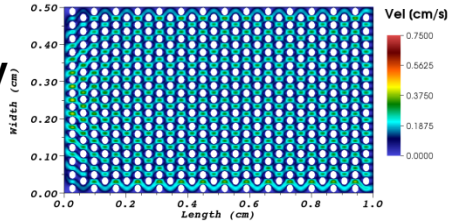
A^{tot} = 4750 m^2 m^{-3}

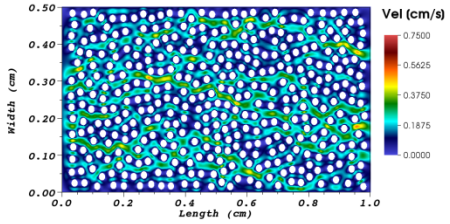
Velocity

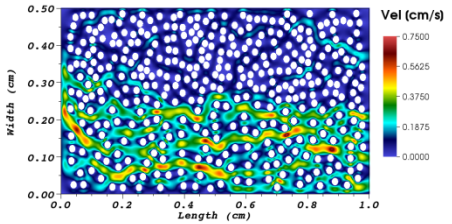
inlet

0.1 cm/s

Regular packing ***Random packing*** ***2-Zone random packing***

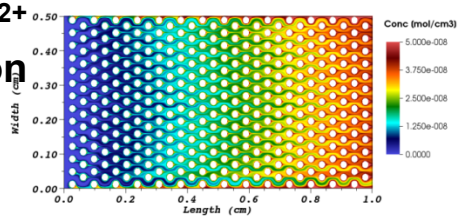


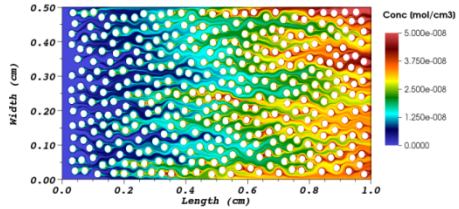


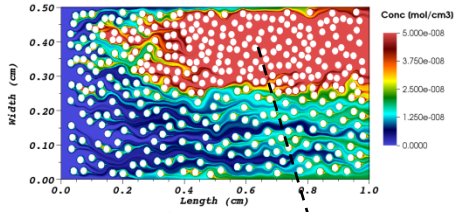


Ca^{2+} Concentration

inlet





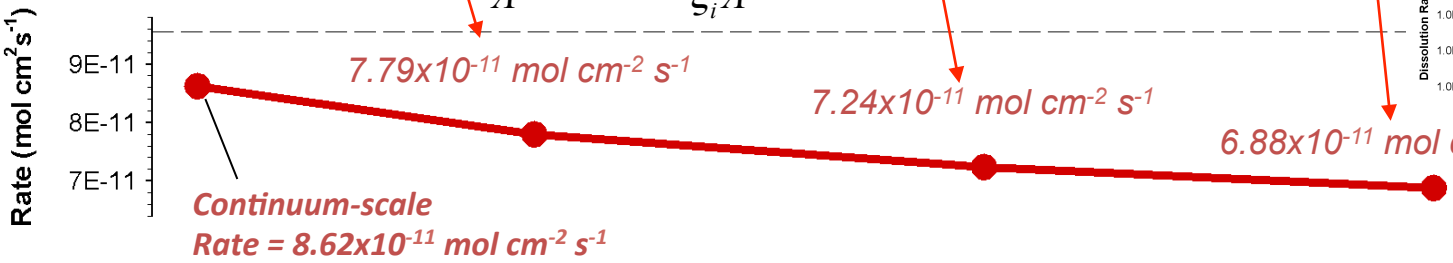


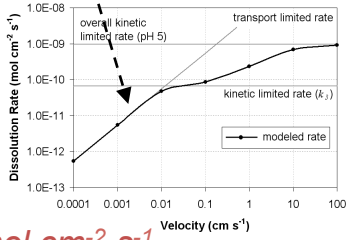
pH 5

0.01 NaCl

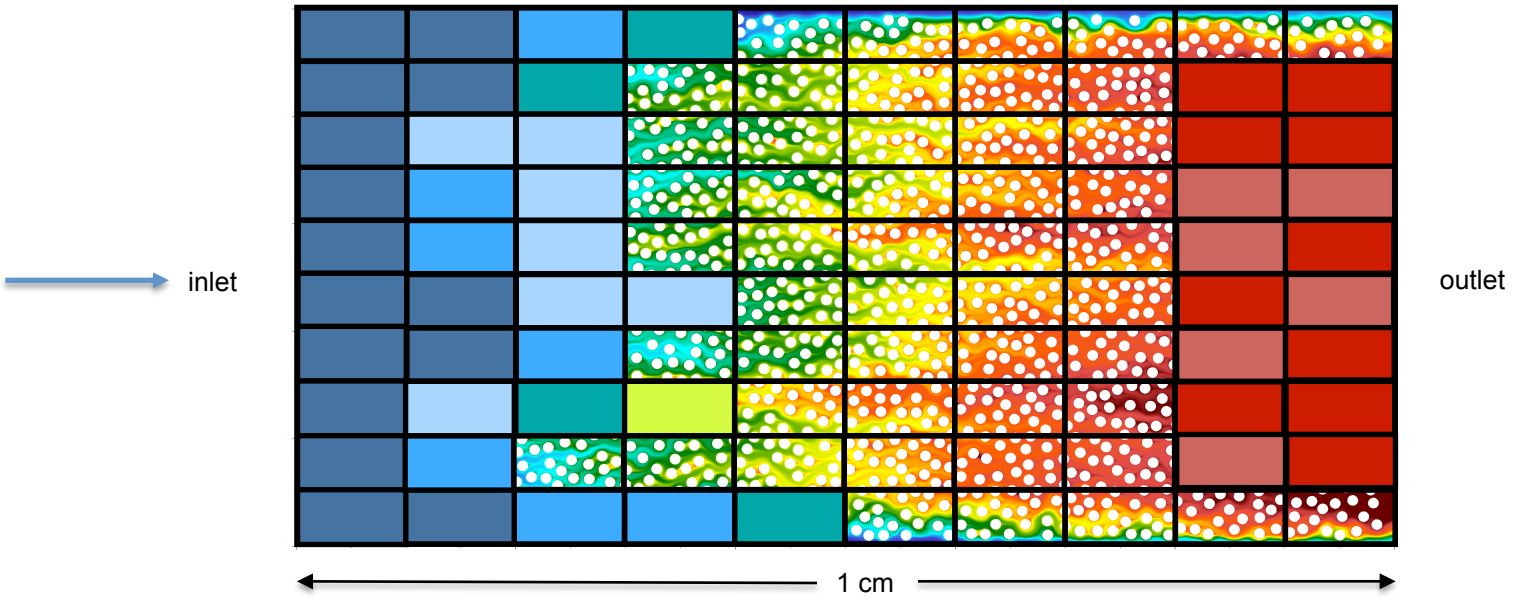
pCO_2 = 3.15 10^{-4} bar

$$\frac{R}{A^{tot}} = \frac{\left(C_i^{outlet} - C_i^{inlet}\right)Q^{tot}}{\xi_i A^{tot}}$$





Volume averaging alone of pore scale does not preserve heterogeneity



$$k = \frac{Q^{tot}\,\mu\,L}{A^{channel}\left(p^{outlet} - p^{inlet}\right)}$$

$$\frac{R}{A^{tot}} = \frac{\left(C_i^{outlet} - C_i^{inlet}\right)Q^{tot}}{\xi_i A^{tot}}$$

We are investigating multiscale methods that preserve heterogeneity of pore scale

Multiscale approach: Use first-of-its-kind highly resolved pore scale simulation data

(i) to inform better parameterizations of permeability, reaction rates and dispersion at continuum Darcy scale, and

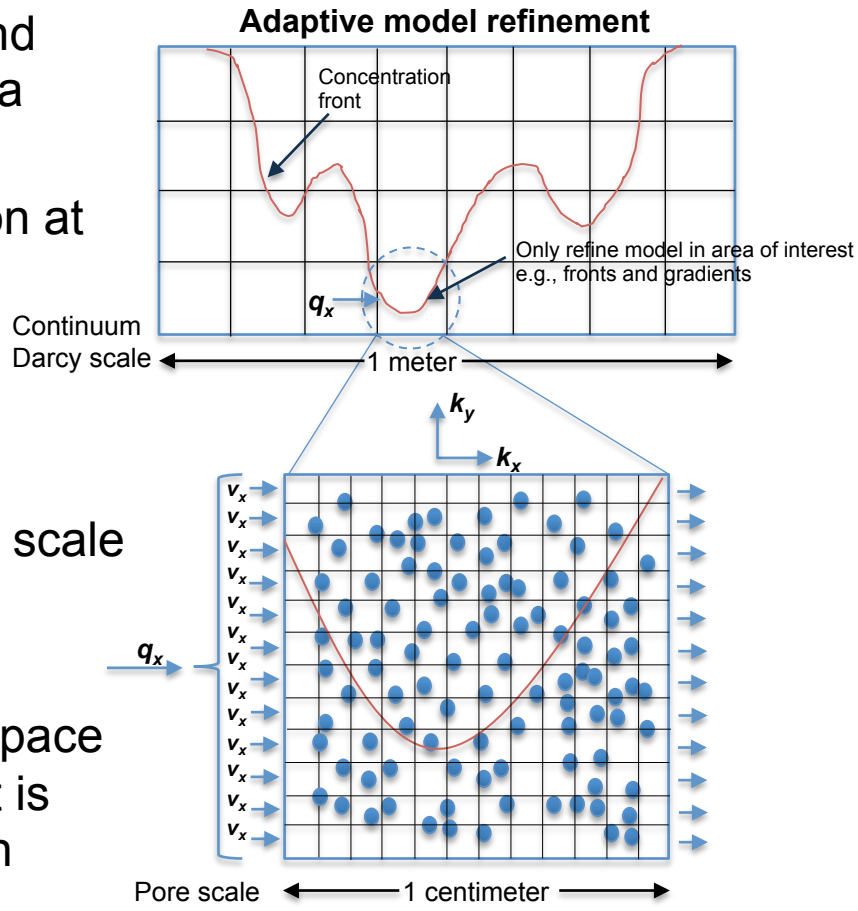
(ii) to verify multiscale approaches:

Deterministic method:

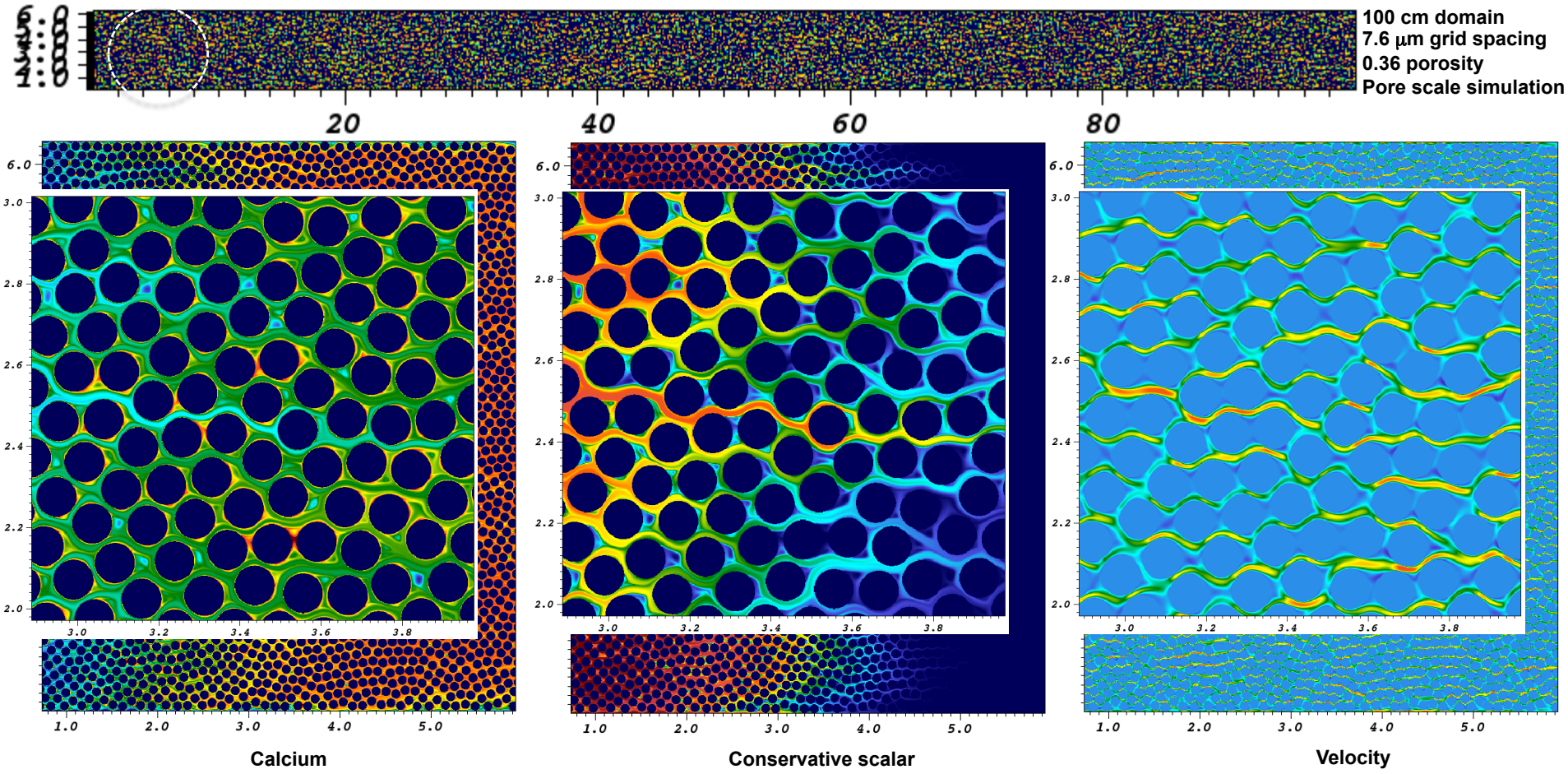
- Adaptive model refinement: directly upscale pore scale data to continuum scale *locally* in areas of interest in domain

Stochastic method:

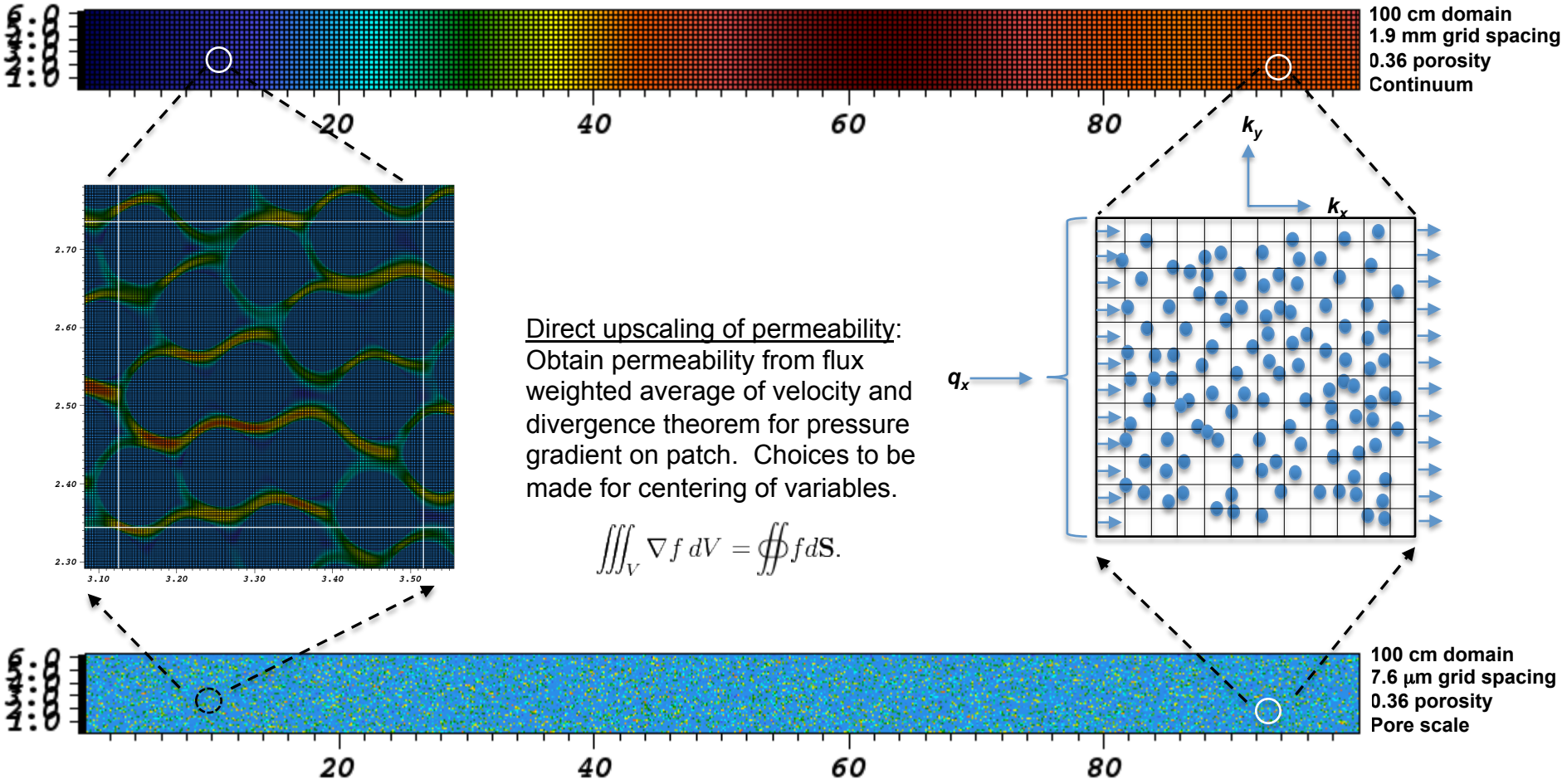
- Intermediate pdf: characterize pore space with a probability density function that is fitted to the pore scale data and graph connectivity



We have established basis for upscaling with high resolution pore scale models



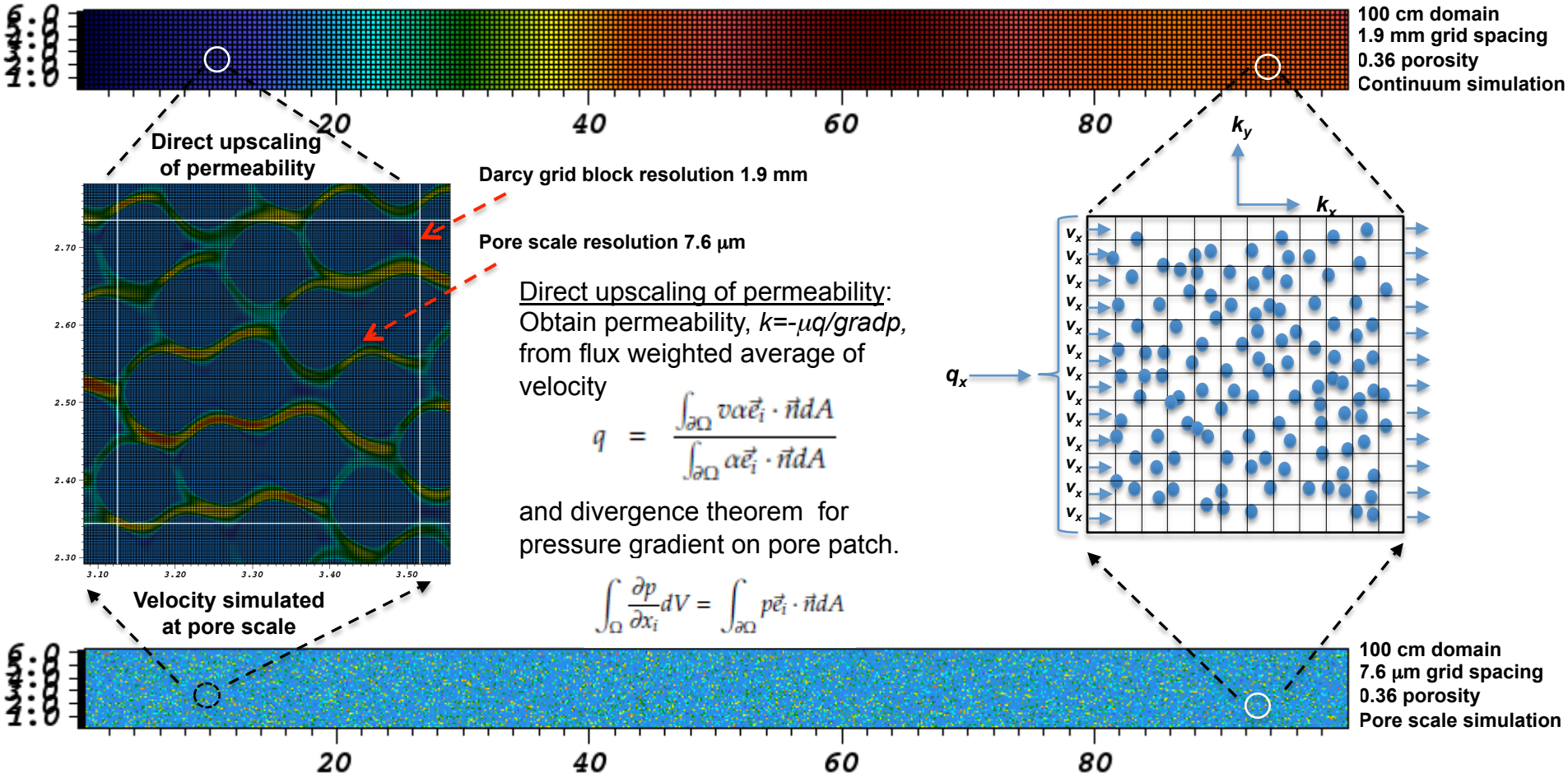
We have constructed algorithms to do direct upscaling from pore to continuum



Direct upscaling of permeability: Obtain permeability from flux weighted average of velocity and divergence theorem for pressure gradient on patch. Choices to be made for centering of variables.

$$\iiint_V \nabla f \, dV = \oiint f d\mathbf{S}.$$

We have constructed algorithms to do direct upscaling from pore to continuum



Direct upscaling of permeability: Obtain permeability, $k=-\mu q/gradp$, from flux weighted average of velocity

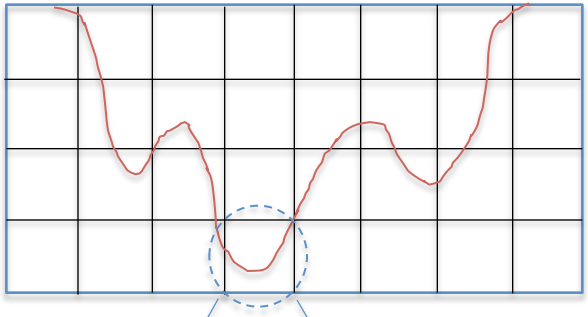
$$q = \frac{\int_{\partial\Omega} v\alpha\vec{e}_i \cdot \vec{n}dA}{\int_{\partial\Omega} \alpha\vec{e}_i \cdot \vec{n}dA}$$

and divergence theorem for pressure gradient on pore patch.

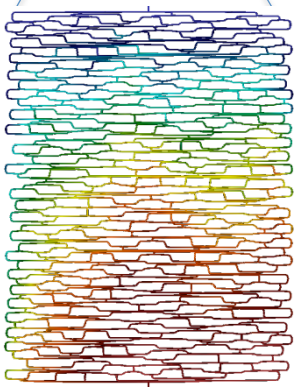
$$\int_{\Omega} \frac{\partial p}{\partial x_i} dV = \int_{\partial\Omega} p\vec{e}_i \cdot \vec{n}dA$$

We can use graph theory to statistically characterize heterogeneous pore space

Continuum Darcy scale



Convert to pore network model (ΔP=RQ) and recouple via Darcy

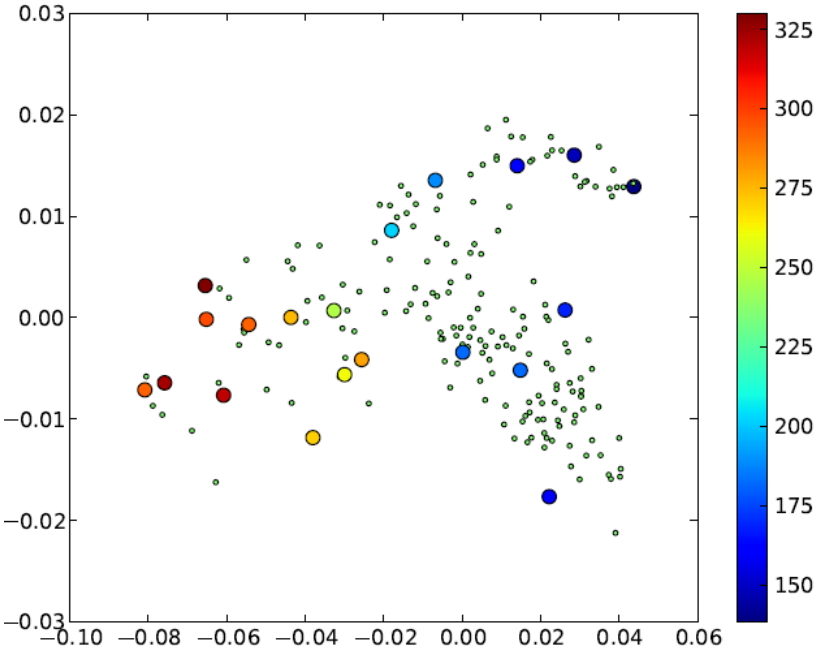


Pore scale simulation with direct upscaling via Darcy

Construct fast pore network solver by using intermediate probability density function to sample characterization of pore space obtained from graph theory and ensemble of pore data

k=-μq/gradp

Single metric derived from graph theory to determine permeability



Principle axes of graph metric using max flow algorithm. 200 packings statistically represented by the metric, tested and compared to 20 simulations (large circles) with dimensionless permeability in color.

Summary

What new science results might be afforded by improvements in NERSC computing hardware, software and services?

- Resolved fluid-fluid, fluid-mineral interactions that shed light on CO_2 trapping mechanisms

Recommendations on NERSC architecture, system configuration and the associated service requirements needed for your science

- A machine with same specs as Edison but with 4-6x the core count

NERSC generally refreshes systems to provide on average a 2X performance increase every year. What significant scientific progress could you achieve over the next 5 years with access to 32X your current NERSC allocation?

- Sub-micron (mesoscale) resolution of reactive transport in shales, carbonates

What "expanded HPC resources" are important for your project?

- Increased core count
- Burst buffer (for both checkpoint resilience and in situ vis and diagnostics)

Chombo-Crunch is an interdisciplinary team of computational and application scientists

Project Team

David Trebotich, PI, LBL-CRD

Carl Steefel, LBL-ESD

Sergi Molins, LBL-ESD

Terry Ligocki, LBL-CRD

Dan Graves, LBL-CRD

Brian Van Straalen, LBL-CRD

Mark Adams, LBL-CRD

Greg Miller, LBL-CRD/UCD

1. Project Description

List of Pi(s)/Institution(s): D. Trebotich (PI), C. Steefel, S. Molins, T. Ligocki, B. Van Straalen, M. Adams, G. Miller; LBNL

- Summarize your project(s) and its scientific objectives through 2017
  - Our goal is to enable accurate prediction of the fate of geologically stored CO_2. We will use high resolution direct numerical simulation to help understand emergent behavior at the pore scale that governs CO_2 trapping mechanisms.
- Our present focus is high resolution single phase reactive transport in carbonates validated by experiments.
- By 2017 we expect to extend this capability to multiple phases and flow and transport in shales.

2. Computational Strategies

- We approach this problem computationally at a high level by using adaptive, finite volume methods based on an embedded boundary approach which enables accurate representation of heterogeneous and microscopic pore space a nd is amenable to moving fluid-mineral and fluid-fluid boundaries.
- The codes we use are based on flow and conservative transport solvers and algorithms developed in the Chombo framework and the multicomponent geochemistry module of CrunchFlow.
- These codes are characterized by these algorithms:
  - Adaptive embedded boundary finite volume methods
  - Algebraic multigrid elliptic solvers (go-through PETSc)
- Our biggest computational challenges are robustness in the presence of non-trivial geometry (cusps, disconnected domains, thin pore throats, etc.)
- Our parallel scaling is limited by memory bandwidth
- We expect our computational approach and/or codes to change (or not) by 2017: multiphase algorithm, moving boundary algorithm, port to Chombo4 (thread-safe, OpenMP)

3. Current HPC Usage (see slide notes)

- Machines currently using (NERSC or elsewhere)
  - Edison, Hopper, Mira
- Hours used in 2012-2013 (list different facilities)
  - ~100M used at NERSC (~40M charged)
- Typical parallel concurrency and run time, number of runs per year
  - 6144, 49,152 for experimental runs; 65,536, 131,072 for benchmark runs
- Data read/written per run
  - 100TB written per 36 hour production run
  - 0.5TB read per run
- Memory used per (node | core | globally)
  - 0.75GB per core
- Necessary software, services or infrastructure
  - PETSc
- Data resources used (/scratch,HPSS, NERSC Global File System, etc.) and amount of data stored
  - /scratch (20TB), HPSS (1-2PB)

4. HPC Requirements for 2017

(Key point is to directly link NERSC requirements to science goals)

- Compute hours needed (in units of Hopper hours)
  - 100M
- Changes to parallel concurrency, run time, number of runs per year
  - Anticipate needing 4x the current core count on Hopper/Edison
- Changes to data read/written
- Changes to memory needed per (core | node | globally)
- Changes to necessary software, services or infrastructure